International Congress on Recent Development in Engineering and Technology (RDET-16) August 22-24, 2016 Kuala Lumpur (Malaysia)

Integrating PCA and GLR for Monitoring Autocorrelated Multivariate Processes

Chun-Chin Hsu[1*] and Fang-Chih Tien[2]

***Abstract*—**Principal Component Analysis (PCA) is a widely used multivariate fault detection method for process monitoring and its counterpart, dynamic PCA (DPCA) was developed when variables exist autocorrelation. Even though DPCA can effectively monitor the autocorrelated processes, the monitoring statistics (T^2 and Q) still exhibit significant autocorrelation which violate the *i.i.d* assumption of the monitoring statistics. Thus, the DPCA with decorrelated residuals (DPCA-DR) was developed to deal with this circumstance. The main disadvantage of DPCA-DR is that it cannot provide practitioner the change point of the process fault which lacks an opportunity for practitioner to diagnose the fault. Based on this point of view, this study strives to integrate DPCA-DR and Generalized Likelihood Ratio (GLR) for monitoring the autocorrelated multivariate processes. Specifically, the DPCA-DR is used for dimension reduction and GLR to be monitoring statistic. The advantages of the proposed method include: 1) it can detect a wide range of process mean shifts, and 2) it provides the practitioner the information of change point which can help diagnose the fault. The efficiency of the proposed method will be verified via a simulated process, in which PCA, DPCA, DPCA-DR will be compared with.

***Index Terms*—**PCA, DPCA, DPCA-DR, GLR

I. Introduction

To ensure product yield and plant safety, immediate fault detection plays an important role for the process production. Statistical Process Control (SPC) is the crucial tool to achieve this goal. The objective of SPC is to online indicate process fault and then find out the root causes and rectify the causes in order to bring the process back to the stably manufactured circumstance. Traditionally, the Shewhart $\overline{X}$ control chart was used to monitor the process change. However, the $\overline{X}$ chart can only detect large process change due to the only consideration of the current observation. Thus, Exponentially Weighted Moving Average (EWMA) and Cumulative Sum (CUSUM) were developed to detect the small process change. Even though, both charts can effectively detect small process change, they cannot effectively detect large process change. In fact, a real process may pose a wide range of process changes. Thus, several researchers endeavor to develop methods to resolve this situation. In general, two main methods were developed to tackle this problem: combined control charts and adaptive control charts. Wu *et al.* [1],[2] proposed the combination of $\overline{X}$ chart and CUSUM chart, in which the $\overline{X}$ chart is used to detect the large process mean shifts and CUSUM used to detect small process changes. However, applying such a way implies the increasing numbers of parameter should be determined by practitioners beforehand. An alternative is to develop adaptive control chart (i.e. adaptive CUSUM and adaptive EWMA charts). This kind of approach firstly estimates the size of any shift that may be occurring, and then uses the estimated value to appropriately adjust control chart in a bid to response effectively to the size of shift. Regarding studies can refer to [3]-[6]. However, the main disadvantage of adaptive chart is that practitioners do not know how to determine control limit at each time. Generalized Likelihood Ratio (GLR) statistic was first presented by R. A. Fisher and was originally developed for hypothesis testing. Recently, the GLR statistic was utilized to develop control chart. The GLR had been shown to be capable of detecting a wide range of process shifts. Besides, unlike several control charts that need practitioners well-tune chart parameters, the GLR chart requires practitioner determine the control limit only. Furthermore, the GLR chart has the advantage of providing diagnostic information, such as estimates of the process change point and estimates of out-of-control mean once an abnormal signal is triggered.

As computer increasingly became available, multivariate measurements for a process can be collected simultaneously. Thus, multivariate SPC (MSPC) was developed for monitor multivariate processes, such as multivariate EWMA, multivariate CUSUM and so forth. Another approach to develop the multivariate monitoring method is the use of data-driven technique. The widely used technique is Principal Component Analysis (PCA). PCA can project the high-dimensional process onto a lower subspace and monitor the process behavior. While the process data exhibits no autocorrelation and stationarity, then PCA can be successfully applied in process monitoring settings with high dimensionality [7]. However, the autocorrelation is commonly seen in a real process data such as health care, semiconductor industry, IT sector and economy. According to [8], the autocorrelation will increase the false alarm rate. Thus, the dynamic PCA (DPCA) developed by [9] was used to monitor the autocorrelation processes. The DPCA introduces the lagged variable for each original variable to augment the original data matrix and then perform traditional PCA for the following analysis. Even though the DPCA can effectively monitor the autocorrelated processes, the calculated monitoring statistics (T^2 and Q) still behavior autocorrelated and this characteristic violates the assumption of *i.i.d* that will lead to an insensitivity to the detection of the fault. Rato and Reis [10] developed DPCA with

Chun-Chin Hsu[1*], Department of Industrial Engineering and Managemnt, Chaoyang University of Technology, Wufong district, Great Taichung, Taiwan.

Fang-Chih Tien[2], Department of Industrial Engineering and Business Managemnt, National Taipei University of Technology, 1, Sec. 3, Zhongxiao E. Rd., Taipei 10608 Taiwan

http://dx.doi.org/10.17758/EAP.EAP816016

decorrelated residuals (DPCA-DR) in a bid to obtain better time-decorrelated statistics. Specifically, DPCA-DR is a combination of DPCA with a one-step-ahead estimation technique.

Even though DPCA-DR is effective to monitor autocorrelation processes, it is insensitive to detect the small process change. Therefore, this study will develop a DPCA-DR-GLR method in an effort to monitor a wide range of process changes. In our proposed methodology, the DPCA-DR is used for the dimension reduction and the GLR is used to monitor the process and provide the estimation of the process change point that give the practitioner the diagnostic information. The efficiency of the proposed method will be verified via a simulated autocorrelated multivariate process, in which the PCA, DPCA and DPCA-DR will be compared. The result demonstrated the proposed method can monitor a wide range of process changes.

II. PCA Related Process Monitoring Methods

A. PCA-based monitoring method

PCA intends to linearly transform high dimensional input vector into a lower dimensional one whose components are uncorrelated. $\mathbf{X}=[\mathbf{x}(1),\mathbf{x}(2),\ldots,\mathbf{x}(n)]\in R^{d\times n}$ denotes a centered data matrix, where $\mathbf{x}\in R^{d}$ is a column vector with *d* measured variables and *n* is the number of measurements. The covariance matrix is $\mathbf{R_x}=E(\mathbf{x}\mathbf{x}^T)$, where *E* represents expectation and *T* denotes the transpose operator. The eigen-decomposition of $\mathbf{R_x}$ can be given by $\mathbf{R_x}=\mathbf{U}\Lambda\mathbf{U}^T$ where $\mathbf{U}\in R^{d\times d}$ is an orthogonal matrix of eigenvectors and $\Lambda\in R^{d\times d}$ is the diagonal matrix of eigenvalues. All score principal components can be expressed as $\mathbf{t}=\mathbf{U}^T\mathbf{x}$. By using only the first few several eigenvectors in descending order of the eigenvalues, the number of principal components in **t** can be reduced, and the reduced **t** is denoted as $\mathbf{t}'\in R^{a}$ where $a\le d$. Denote $\mathbf{P}\in R^{d\times a}$ and $\mathbf{D}\in R^{a\times a}$ as the matrices of eigenvectors and eigenvalues, respectively. They are associated with the retained principal components such that $\mathbf{t}'=\mathbf{P}^T\mathbf{x}$. Hotelling's T^2 can be used to measure the variation of systematic part of PCA model. T^2 is the sum of the normalized squared scores, that is, $T^2=\mathbf{t}'^T\mathbf{D}^{-1}\mathbf{t}'=\mathbf{x}^T\mathbf{P}\mathbf{D}^{-1}\mathbf{P}^T\mathbf{x}$. The upper confidence limit for T^2 can be obtained by using *F* distribution, and it takes the form as $T^2_{a,n,\alpha}=\frac{a(n-1)}{n-a}F_{a,n-a,\alpha}$. A measure of variation not captured by the PCA model can be monitored by the Squared Prediction Error (*SPE*). $SPE=\mathbf{e}^T\mathbf{e}=\mathbf{x}^T(\mathbf{I}-\mathbf{P}\mathbf{P}^T)\mathbf{x}$, where residual is $\mathbf{e}=\mathbf{x}-\hat{\mathbf{x}}=\mathbf{x}-\mathbf{P}\mathbf{t}'=(\mathbf{I}-\mathbf{P}\mathbf{P}^\mathbf{T})\mathbf{x}$ and $\mathbf{e}=0$ if $a=d$. The upper control limit for *SPE* can be expressed as $SPE_\alpha=\theta_1\left[\frac{c_\alpha\sqrt{2\theta_2h_0^2}}{\theta_1}+1+\frac{\theta_2h_0(h_0-1)}{\theta_1^2}\right]^{1/h_0}$. where $\theta_g=\sum_{j=a+1}^{d}\lambda_j^g$ for $g=1,2,3$, $h_0=1-\frac{2\theta_1\theta_3}{3\theta_2^2}$ and c_α is the normal deviate corresponding to upper $1-\alpha$ percentile.

B. DPCA-based monitoring method

Dynamic PCA (DPCA) was first introduced by [9] as a way to monitor the multivariate autocorrelated processes. In DPCA, the originally collected data matrix is augmented by the added lagged variables. The lagged variables means that in addition to the current observed variables, the respective lagged values up to a proper order *l*, can also be included, that is:

$$\mathbf{X}(l)=\begin{matrix}\\ \text{Sample }1\\ \text{Sample }2\\ \vdots\\ \text{Sample }n\end{matrix}\begin{matrix} lag0 & lag1 & \cdots & lagl\\ \left[\begin{array}{c|c|c|c}\mathbf{x}_1^T & \mathbf{x}_0^T & \cdots & \mathbf{x}_{1-l}^T\\ \mathbf{x}_2^T & \mathbf{x}_1^T & \cdots & \mathbf{x}_{2-l}^T\\ \vdots & \vdots & \vdots & \vdots\\ \mathbf{x}_n^T & \mathbf{x}_{n-1}^T & \cdots & \mathbf{x}_{n-l}^T\end{array}\right]\end{matrix}\in\mathbf{R}^{n\times dl}$$

After augmenting the original data matrix, the traditional PCA method can be performed to the augmented data matrix.

C. DPCA-DR-based monitoring method

Even though the DPCA is effective to deal with autocorrelated process, its monitoring statistics still exhibit autocorrelation. To deal with this problem, Rato and Reis [10] developed DPCA with decorrelated residuals (DPCA-DR). Specifically, DPCA-DR is a combination of DPCA with a missing data estimation technique in order to obtain better time-decorrelated statistics. The DPCA-DR firstly obtain the scores from the implementation of DPCA, that is, $y_a=P_a^{'}(x-\bar{x})$. An additional vector of estimated scores $\hat{y}_a$ is computed by assuming the current observation vector is missing. This is a one-step-ahead prediction of the scores based on implicit AR model estimated by DPCA. The Hotelling T^2 is defined: $T^2_{prev}=(y_a-\hat{y}_a)^{'}S^{-1}_{prev}(y_a-\hat{y}_a)$, where S_{prev} denotes the sample covariance matrix of the difference between the observed and estimated score, $(y_a-\hat{y}_a)$. The *SPE* is defined as $SPE_{res}=(x-P_ay_a)^{'}S^{-1}_{res}(x-P_ay_a)$, where S_{res} is the sample covariance matrix of the residuals in the reconstructed data, obtained with the estimated scores.

D. GLR-based monitoring

Considering data $\{X_1,X_2,\cdots,X_k\}$ were *iid* drawn from the process and the distribution follows $N(\mu_0,\sigma_0^2)$, where the suffix denotes the sample time. The likelihood function at time *k* can be represented as: $L(\infty,\mu_0|X_1,X_2,\cdots,X_k)=(2\pi\sigma_0^2)^{-k/2}\exp\left(-\frac{1}{2\sigma_0^2}\sum_{i=1}^{k}(X_i-\mu_0)^2\right)$. If the process mean shifts to some value $\mu_1\ne\mu_0$ at some time τ^* between τ and $\tau+1$. The likelihood function at sample *k* can be represented as: $L(\infty,\mu_1|X_1,X_2,\cdots,X_k)=(2\pi\sigma_0^2)^{-k/2}\exp\left(-\frac{1}{2\sigma_0^2}\left(\sum_{i=1}^{\tau}(X_i-\mu_0)^2+\sum_{i=\tau+1}^{k}(X_i-\mu_1)^2\right)\right)$. The log likelihood ratio statistic at sample *k* can be exhibited as

$$R_k = \ln \frac{\max\limits_{0\le\tau\le k} L(\tau,\mu_1 \mid X_1, X_2, \cdots X_k)}{L(\infty,\mu_0 \mid X_1, X_2, \cdots X_k)} = \max_{0\le\tau\le k} \frac{(\hat{\mu}_{1,\tau,k}-\mu_0)}{\sigma_0^2}\left\{\sum_{i=\tau+1}^{k}\left[(X_i-\mu_0)-\frac{1}{2}(\hat{\mu}_{1,\tau,k}-\mu_0)\right]\right\}$$

Where $\hat{\mu}_{1,\tau,k}$ is the maximum likelihood estimator (MLE) of μ_1, that is $\hat{\mu}_{1,\tau,k} = \frac{\sum_{i=\tau+1}^{k} X_i}{k-\tau}$. Therefore, R_k can be reduced to $R_k = \max_{0\le\tau\le k} \frac{(k-\tau)}{2\sigma_0^2}(\hat{\mu}_{1,\tau,k}-\mu_0)^2$. To reduce the computational complexity, only the recent past *m* samples (i.e. window size) are used to seek the maximum value [11]. Let $\hat{\tau}$ be the estimation of the process change-point at which the maximum value has been reached. The GLR statistic can been then expressed as: $R_{m,k} = \max_{0\le\tau\le k} \frac{(k-\hat{\tau})}{2\sigma_0^2}(\hat{\mu}_{1,\tau,k}-\mu_0)^2$. A signal is given at sample *k* if $R_{m,k} > h_{GLR}$. Another form for expressing GLR chart is to calculate $R'_{m,k}$, that is: $R'_{m,k} = \sqrt{2R_{m,k}}\sigma_0 = \sqrt{k-\hat{\tau}}(\hat{\mu}_{1,\hat{\tau},k}-\mu_0)$. If $R'_{m,k}$ falls outside the interval of $[-h'_{GLR}\sigma_0, h'_{GLR}\sigma_0] = [-\sqrt{2h_{GLR}}\sigma_0, \sqrt{2h_{GLR}}\sigma_0]$, then a signal will be triggered.

III. Proposed Monitoring Method

This study will develop a monitoring methodology in an attempt to detect a wide range of the process changes for an autocorrelated multivariate process. Specifically, the proposed method integrated DPCA-DR and GLR, in which, the DPCA-DR is used to reduce the dimensionality and the GLR is utilized as a monitoring statistic. For short, the proposed methods is called DPCA-DR-GLR. Figure 1 shows the flow chat of DPCA-DR-GLR. In general, the proposed method consist of two main phases, Phase I for the offline training and Phase II for the online fault detection. The Phase I using normal data to construct monitoring model and estimate relevant parameters. For Phase II, the new data collected from online process is used to judge the process status. Details are specified as below:

Phase I: Offline training

Step 1: Obtaining an uncontaminated dataset $X_{normal} \in R^{n\times d}$ from a normal-operated process, where *n* denotes the number of observations and *d* denotes the dimension.

Step 2: Calculating KSV (Key Singular Values) and KSVR (Ratio of Successive Key Singular Values) for selecting the number of lags for each original variable $l = (l_1, l_2, \cdots, l_d)$. The selected lagged structure generates the lowest KSV and KSVR.

Step 3: Scaling the augmented data matrix by sample mean ($\hat{\mu}_0$) and standard deviation ($\hat{\sigma}_0$), the scaled data matrix denotes as $\tilde{X}_{normal}(l)$.

Step 4: Performing PCA and we can obtain the eigenvector $\mathbf{U} \in R^{d\times d}$ and its eigenvalues $\Lambda \in R^{d\times d}$. By using Cumulative Percentage of Variance (CPV), the dimension reduction can be achieved. The CPV is a measure of how much variation is captured by the first *a* latent variables:

$$CPV = \frac{\sum_{j=1}^{a} u_j}{\sum_{j=1}^{d} u_j} \times 100\%$$

In this study, the value of CPV is set to be 95%. After determining the number of retained components, its corresponding eigenvectors and eigenvalues are $\mathbf{P_a} \in R^{d\times a}$ and $\mathbf{D_a} \in R^{a\times a}$. Thus, the retained components can be obtained from $y_a = P'_a \tilde{X}_{normal}(l)$.

Step 5: Performing the one-step-ahead prediction of the components based on AR model and it is detonated as $\hat{y}_a$. Calculating $y_a - \hat{y}_a$ and its covariance matrix S_{prev}, then the Hotelling statistic can be obtained by $T^2_{prev} = (y_a - \hat{y}_a)' S^{-1}_{prev}(y_a - \hat{y}_a)$.

Step 6: Estimating the mean ($\hat{\mu}_{T^2_{prev}}$) and standard deviation ($\hat{\sigma}_{T^2_{prev}}$) based on T^2_{prev}.

Step 7: Setting the window size to be 400 as proposed by [12] and calculate the charting statistic R'.

Step 8: According to the research from [12], they found there is a linear relationship between h_{GLR} (i.e. the control limit for charting statistic R) and the log scale of false alarm rats as shown in Figure 2. The equation can be expressed as

$$h_{GLR} \approx -0.87 + 1.12\ln(\textit{false alarm rate} \mid m = 400).$$

For the charting statistic R', the upper control limit and lower control limit can be acquired by

$$UCL_{R'} = \sqrt{2h_{GLR}}\sigma_{T^2_{prev}}$$

$$LCL_{R'} = -\sqrt{2h_{GLR}}\sigma_{T^2_{prev}}$$

Phase II: Online monitoring

Step 1: Online obtain a new dataset $X_{new} \in R^{n\times d}$

Step 2: Using $l = (l_1, l_2, \cdots, l_d)$ to augment the data matrix.

Step 3: Scaling the data matrix as the same step in Phase I.

Step 4: Obtaining the Hotelling statistic $T^2_{prev,new}$.

Step 5: Calculating the charting statistic R'_{new}

Step 6: If $R'_{new} > UCL_{R'}$ or $R'_{new} < LCL_{R'}$ then it indicates an abnormal situation exists in the process and the practitioner should immediately find out the root causes and rectify it. The estimation of the change point (τ) provide practitioner the diagnostic information, such as to trace back which production batch maybe the start of the faulty batch.

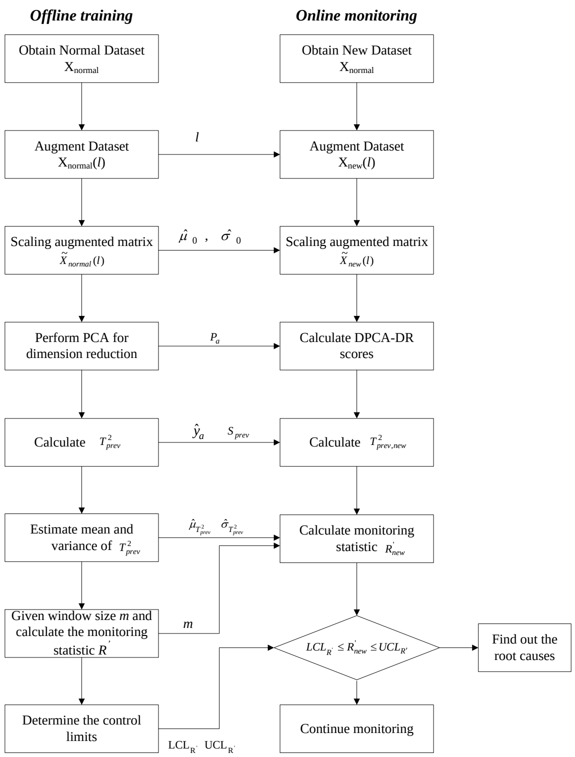

Fig. 1. DPCA-DR-GLR methodology

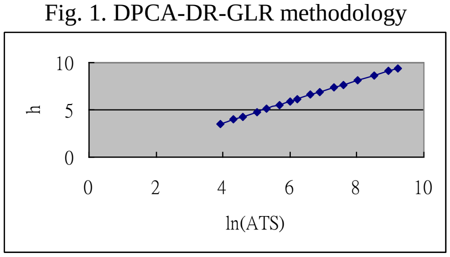

Fig. 2 h_{GLR} v.s. the log scale of false alarm rate

IV. A SIMULATED PROCESS

Suppose a simulated multivariate non-Gaussian process, suggested by [9] and is given as

$$\mathbf{r}(k)=\begin{bmatrix}0.118 & -0.191 & 0.287\\ 0.847 & 0.264 & 0.943\\ -0.333 & 0.514 & -0.217\end{bmatrix}\mathbf{r}(k-1)+\begin{bmatrix}1 & 2\\ 3 & -4\\ -2 & 1\end{bmatrix}\mathbf{u}(k-1)$$

$$\mathbf{g}(k)=\mathbf{r}(k)+\mathbf{v}(k)$$

where $\mathbf{g}$ is the output, $\mathbf{r}$ is the state and $\mathbf{v}$ denotes the input which is assumed to be normal distributed with zero mean and variance of 0.1. The input $\mathbf{u}$ is given by

$$\mathbf{u}(k)=\begin{bmatrix}0.811 & -0.226\\ 0.477 & 0.415\end{bmatrix}\mathbf{u}(k-1)+\begin{bmatrix}0.193 & 0.689\\ -0.320 & -0.749\end{bmatrix}\mathbf{h}(k-1)$$

The input $\mathbf{h}$ is assumed to be uniformly distributed with random vector over interval (-2,2). The five variables $(g_1, g_2, g_3, u_1, u_2)$ are used to monitor the process. Normal data with 500 samples are used for the analysis. The faults for monitoring and diagnosis are as follow:

Fault 1 (large shift): a step change of h_1 by 3 is introduced at sample 200 in 500 simulated runs

Fault 2 (small shift): a step change of h_2 by 1.5 is introduced at sample 200 in 500 simulated runs

For a fair comparison, all methods used the empirical limits with false detection rate of 1%. Furthermore, the CPV is used to select the number of dominant components (i.e. CPV>95%). Figure 3 shows the PCA monitored result, in which the number of component retained is 3. The figure indicates PCA cannot effectively detect the fault 1 due to the autocorrelation characteristic in the process.

For DPCA implementation, Figure 4 shows the KSV (Key Singular Values) and KSVR (Ratio of Successive Key Singular Values) for selecting the number of lags for each original variable. It indicates an optimum lagged structure at stage 4. Thus, the number of the lags used for g_1, g_2, g_3, u_1, u_2 are 0,0,1,2,1, respectively. Figure 5 demonstrates the monitored result and it also cannot effectively detect the fault because of the autocorrelated monitoring statistics as shown in Figure 6.

Figure 7 shows the DPCA-DR monitoring result of fault 1, it indicates *SPE* can detect the fault after sample 200. However, the T^2 still cannot detect the fault. The Figure 8 shows the autocorrelation function for the monitored *SPE* and it indicates that the less autocorrelation in the monitored statistic that prompts a better monitoring result than previous mentioned monitored methods.

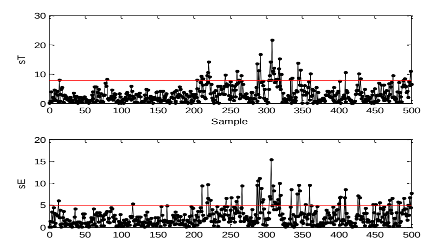

Fig. 3. PCA monitored result for fault 1

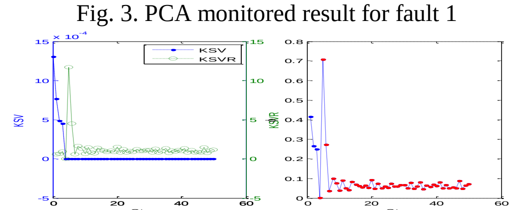

Fig. 4. Selection of the number of lags for DPCA

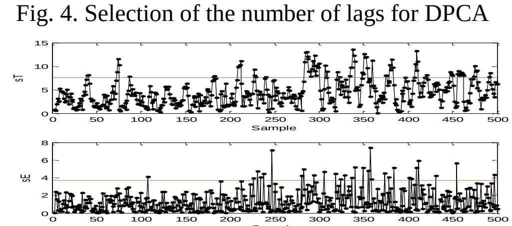

Fig. 5. DPCA monitored result for fault 1

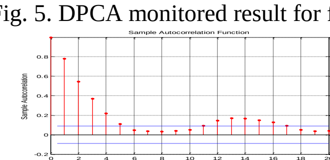

Fig. 6. Autocorrelation function for monitored T^2

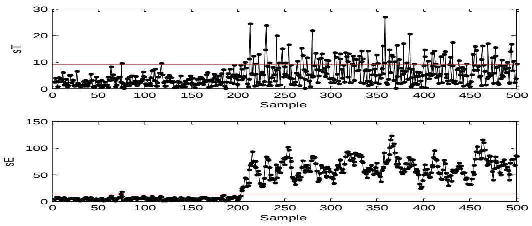

Fig. 7. DPCA-DR monitored result for fault 1

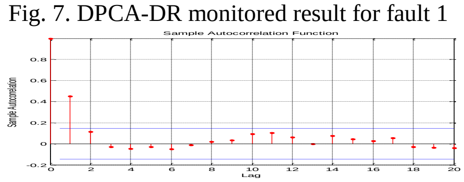

Fig. 8. Autocorrelation function for monitored *SPE*

For implementation of the proposed method, the window sizes is set to be 400 for the GLR implementation. The estimated mean and standard deviation for T^2_{prev} under normal data are $\hat{\mu}_{T^2_{pre}} = 2.9105$ and $\hat{\sigma}_{T^2_{pre}} = 1.8767$. The control limit for $R_{m,k}$ is: $h_{GLR} = -0.87 + 1.12\ln(\textit{false alarm rate} = 100 \mid m = 400) = 4.28$ and control limit of $R^{'}_{m,k}$ is $\left[-\sqrt{2h_{GLR}}\hat{\sigma}_{0,T^2_{pre}}, \sqrt{2h_{GLR}}\hat{\sigma}_{0,T^2_{pre}}\right] = [-8.51, 8.51]$. Figure 9 shows the monitor result of the proposed method, it indicated that the GLR chart can detect the fault immediately. Figure 10 is the estimation of the change point and the plot shows the convergence at 200. The figure 9 also exhibits an upward trend around sample 200. This change point information can help practitioner to diagnose the fault.

In the case of small process change (Fault 2), Figure 11 and Figure 12 demonstrate the monitoring results of DPCA-DR and the proposed method, respectively. It shows the DPCA-DR cannot detect the small process change, whereas our proposed method can effectively indicate the occurrence of the fault.

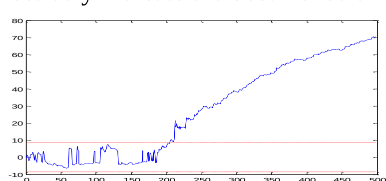

Fig. 9. The monitoring result of the proposed method (DPCA-DR-GLR) for fault 1

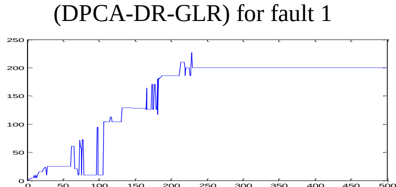

Fig. 10. The plot of the estimated change point

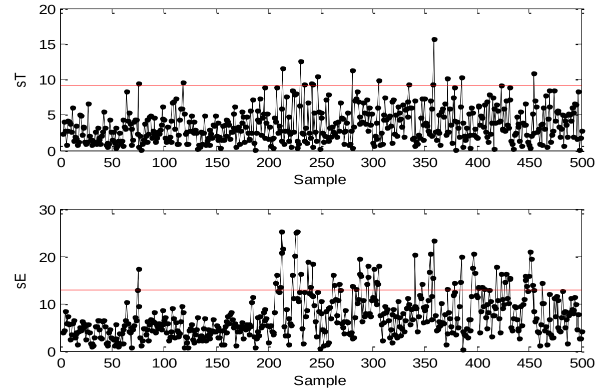

Fig. 11. DPCA-DR monitored result for fault 2

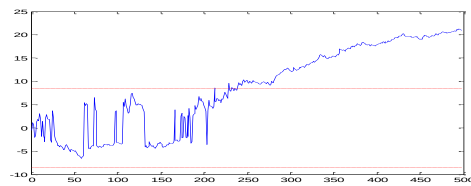

Fig.12. The monitoring result of the proposed method (DPCA-DR-GLR) for fault 2

V. Conclusion

Due to IT increasingly became available, the high-dimensional data can be easily collected online and the frequent sampling of the data will lead to autocorrelation exhibiting in the data structure. Thus, developing multivariate autocorrelated process monitoring method is a crucial work for the modern industry. In this study, we proposed a DPCA-DR-GLR method for detecting a wide range of process changes. The proposed methods utilized the DPCA-DR as the dimension reduction tool, whereas the GLR used as the monitoring tool to detect the process fault. According to the implementation result, it indicated the PCA unable to detect autocorrelated process fault due to PCA only deliberates on the current observation. The DPCA also performed poorly because of the autocorrelated monitoring statistics. For DPCA-DR, it can detect the autocorrelated processes well because of the decorrelated monitoring statistics. However, the DPCA-DR is insensitive to the small process change. Conversely, the result shows the proposed method can detect both the large and small process changes and further provide the estimation of the change point that will help the practitioner diagnose the process fault. For the future study, researcher can put efforts on the non-Gaussian process monitoring since the PCA has the limit of the Gaussianity assumption.

References

[1] Wu, Z.; Yang, M.; Jiang, W.; and Khoo, M. B. C. (2008). "Optimization Designs of the Shewhart-CUSUM Control Charts". *Computational Statistics and Data Analysis* 53, pp. 496-506. http://dx.doi.org/10.1016/j.csda.2008.08.032

[2] Wu, Z.; Yang, M.; Khoo, M. B. C.; and Yu, F. J. (2010). "Optimization Designs and Performance Comparisons of Two CUSUM Schemes for Monitoring Process Shifts in Mean and Variance". *European Journal of Operational Research* 205, pp. 136-150. http://dx.doi.org/10.1016/j.ejor.2009.12.005

[3] Reynolds, M. R., JR. and Stoumbos, Z. G. (2006). “Comparisons of Some Exponentially Weighted Moving Average Control Charts for Monitoring the Process Mean and Variance”. *Technometrics* 48, pp. 550-567. http://dx.doi.org/10.1198/004017006000000255

[4] Shu, L. J. and Jiang, W. (2006). “A Markov Chain Model for the Adaptive CUSUM Control Chart”. *Journal of Quality Technology* 38, pp. 135-147.

[5] Jiang, W.; Shu, L.; and Apley, D. W. (2008). “Adaptive CUSUM Procedures with EWMA-Based Shift Estimators”. *IIE Transactions* 40, pp. 992-1003. http://dx.doi.org/10.1080/07408170801961412

[6] Wu, Z.; Jiao, J.; Yang, M.; Liu, Y.; and Wang, Z. (2009). “An Enhanced Adaptive CUSUM Control Chart”. *IIE Transactions* 41, pp. 642-653. http://dx.doi.org/10.1080/07408170802712582

[7] De Ketelaere, B., Hubert, M., and Schmitt, E. (2015). ”Overview of PCA-based statistical process monitoring methods for time-dependent, high-dimensional data”. *Journal of Quality Technology* 47(4), pp. 318-335.

[8] Hwarng, H. B., Wanng, Y. (2010). Shift detection and source identification in multivariate autocorrelated process. International Journal of Production Research 48(3), pp. 835-859. http://dx.doi.org/10.1080/00207540802431326

[9] Ku, W., Storer, R. H., & Georgakis, C. (1995). Disturbance detection and isolation by dynamic principal component analysis. *Chemometrics and Intelligent Laboratory Systems*, 30, 179-196. http://dx.doi.org/10.1016/0169-7439(95)00076-3

[10] Tiago, J. Rato and Marco, S. Reis. (2013). Fault detection in the Tennessee Eastman benchmark process using dynamic principal components analysis based on decorrelated residuals (DPCA-DR). *Chemometrics and Intelligent Laboratory Systems*, 125, 101-108 http://dx.doi.org/10.1016/j.chemolab.2013.04.002

[11] A. S. Willsky and H. L. Jones, A generalized likelihood ratio approach to detection and estimation of jumps in linear systems, *IEEE Trans. on Automatic Control*, vol.21, pp.108-112, 1976. http://dx.doi.org/10.1109/TAC.1976.1101146

[12] Reynolds, M. R., Jr. and Lou, J. (2010). “An Evaluation of a GLR Control Chart for Monitoring the Process Mean”. *Journal of Quality Technology* 42, pp. 287-310.